

Revisiting Cohesive Devices in Academic L2 English Writing:
What DO Successful Writers Use?

Castle Sinicrope
SLS-672: Classroom Research
Dr. McKay
April 25, 2007

PART ONE: INTRODUCTION

In the thirty years of research and discussion following the publication of Halliday and Hasan's seminal work *Cohesion in English* (1976), no definitive answers have emerged regarding the relationship of cohesive devices and quality of writing. Simply stated, cohesion refers to the set of linguistic features for creating a text out of sentences; cohesion is a means for combining sequences of sentences together to form an integrated whole. For example, the cohesive device *instead* links the following two sentences together as a unit, creating a relationship of contrast between the first and the second sentence: "He showed no pleasure at hearing the news. *Instead* he looked even gloomier" (Halliday and Hasan, 1976, p. 254). Although researchers have studied the use of cohesive devices across a variety of contexts, native and non-native writing in foreign and second language settings, the findings remain contradictory, and the debate over role of cohesive devices in good writing continues.

Although some researchers have found a correlation between cohesive devices and writing quality (Liu and Braine, 2005; Wenjun, 1999; Witte and Faigley, 1981), others have shown no difference in the use of cohesive devices in "good" and "weak" writing. (Johnson, 1992; Tierney and Mosenthal, 1983; Zhang, 2000). Another group of researchers falls between the two poles, showing difference for certain types of cohesion (Neuner, 1987; Hinkel, 2001; Yang, 1989). To date, however, no studies have provided an in-depth description of the types and uses of cohesive devices in successful L2 writing, particularly across different academic genres. To do so, this study examines the types and frequencies of cohesive devices in the writing of four successful L2 writers who have completed writing tasks that represent three different academic genres: definition, analysis, and research papers. By revisiting the question of cohesion and approaching the question from a new angle, this pilot study seeks to map a

landscape of cohesion in what is perceived as “successful” L2 writing and provide pedagogical suggestions for the classroom.

1.1 Theoretical framework

For past and contemporary researchers alike, *Cohesion in English* (Halliday and Hasan, 1976) has provided an important framework for identifying and analyzing cohesive devices in spoken and written discourse. In their work, Halliday and Hasan specify five types of cohesion: reference, ellipsis, substitution, conjunction, and lexical cohesion. Each cohesive device type consists of several subtypes.

The first type of cohesive device, reference, contributes to cohesion by pointing to another element in the text, thereby creating a continuity of reference. Under the broad type *reference* Halliday and Hasan differentiate between three subtypes: pronomial reference, demonstratives (including the definite article), and comparatives. All of these subtypes may be anaphoric or cataphoric in their reference, i.e., the referent may proceed or follow the reference device in the text. In (1), the personal pronoun *they* establishes a cohesive link between the first and the second sentence by pointing back to the nominal phrase *three blind mice*. Similarly, the demonstrative *these* in (2) points back to the favored lions and polar bears in the previous sentence and the definite article *the* in (3) refers to the previously mentioned hall. Finally, the comparative tie in (4), *more things*, creates cohesion by pointing back to the aforementioned wrens. For a complete list of reference devices and examples, see Appendix A.

- (1) ***Three blind mice, three blind mice.***
See how ***they*** run! See how ***they*** run! (Ibid, p. 31)
- (2) I like ***the lions***, and I like ***the polar bears***. ***These*** are my favorites (Ibid, p. 60).
- (3) She found herself in ***a long, low hall***, which was lit up by a row of lamps hanging from the roof. There were doors all around ***the hall***, but they were all locked (Ibid, p. 72).

- (4) There were *two wrens* upon a tree.
Another came, and there were three (Ibid, p. 31).

While reference typifies cohesion on the semantic level, the second and third types, substitution and ellipsis, represent grammatically cohesive relations. In substitution, one item is replaced with another; in ellipsis, an item is omitted. Example (5) illustrates the substitution of *does* for *knows*, and example (6) shows the omission of the verb *brought* in the latter half of the sentence (Ibid, Chapters 3 and 4).

- (5) You think Joan already *knows*? – I think everybody *does* (Ibid, p. 91).
(6) Joan *brought* some carnations, and Catherine some sweet peas (Ibid, p. 143).

None of the studies of cohesion in writing use the categories of substitution and ellipsis to analyze textual cohesion, and several studies provide rationale for excluding these two types (Liu and Braine, 2005; Tierney and Mosenthal, 1983; Wenjun, 1999; Witte and Faigley, 1981). Reasons for exclusion include zero or near-zero occurrence in pilot studies and the fact that substitution and ellipsis are characteristic of spoken not written discourse.

Conjunction, the fourth type, contributes to textual cohesion by linking successive sentences that are not structurally related. Under the broad category of conjunction, four subcategories capture different relations: additive, adversative, causal, and temporal conjunctions. Additive conjunctives, as demonstrated in (7), express addition or similarity between elements (e.g., *and, in addition to, furthermore*). In contrast, adversative conjunctives create an alternative, dissimilar, or unexpected relation between successive elements, as in (8) (e.g., *however, on the other hand, instead*), or set the stage for exemplification or apposition (e.g., *that is to say, in other words*). Causal conjunctives, illustrated in (9), build a relation of result, reason, or purpose between elements (e.g., *for this reason, as a result of, to this end*) and provide a context for emphatic and respective relations (e.g., *in that case, with regard to this*). Finally, temporal

conjunctives order sentences in a text sequentially, whether preceding, following, or occurring simultaneously with related elements (e.g., *previously*, *subsequently*, *simultaneously*, *first*, *lastly*, *in conclusion*) (Halliday and Hasan, 1976, Chapter 5). See Appendix B for a full list of conjunctives and their subtypes.

For the whole day he climbed up the steep mountainside, almost without stopping

- (7) **And** in all this time he met no one (additive).
- (8) **Yet** he was hardly aware of being tired (adversative).
- (9) **So** by night time the valley was far below him (causal).
- (10) **Then**, as dusk fell, he sat down to rest (temporal). (Ibid pp. 238-239)

The fifth and final type of cohesion, lexical cohesion, encompasses five subtypes that create cohesion through reiterative choices in the lexis: repetition, synonym, superordinate, general nouns, and collocations. In repetition, cohesion is achieved through repeating the same word or phrase (11). Synonyms and superordinates establish cohesive ties between elements by pointing to the original referent with a different lexical form while expressing the same or expanded semantic meaning (12), (13). General nouns, such as *thing* or *person*, exist on the boundaries of lexical cohesion and substitution. In (14), the general noun *thing* functions as lexical device by referring back to the nominal phrase *the ascent of the peak*. The last lexically cohesive device, collocation, achieves cohesion in texts through commonly co-occurring lexical items. Because collocation presents a variety of issues, Halliday and Hasan label it the “most problematical” aspect of lexical cohesion (p. 284). No simple and short examples can illustrate collocation’s cohesive function (Ibid, Chapter 6).

I turned to the *ascent of the peak*.

- (11) The **ascent** is perfectly easy (repetition).
- (12) The **climb** is perfectly easy (synonym).
- (13) The **task** is perfectly easy (superordinate).
- (14) The **thing** is perfectly easy (general word). (Ibid, p. 279)

1.2 Past Research

Studies on the relationship between cohesive devices and writing quality have yielded contradictory results over the past thirty years. Although researchers have applied the same framework as proposed by Halliday and Hasan, findings range from a strong correlation between use of cohesive devices and writing quality (Liu and Braine, 2005; Wenjun, 1999; Witte and Faigley, 1981) to no correlation between the two (Johnson, 1992; Tierney and Mosenthal, 1983; Zhang, 2000). Both groups of studies include participants from the same groups: L1 Chinese writers of English (Liu and Braine, 2005; Wenjun, 1999; Zhang, 2000) and L1 English writers using their first language (Johnson, 1992; Tierney and Mosenthal, 1983; Witte and Faigley, 1981). To further complicate the existing body of research, some studies show correlation for some types of cohesive devices but not others (Liu and Braine, 2005, Johnson, 1992; Neuner, 1987; Wenjun, 1999). Whether described in terms of statistical correlations or frequency counts, the relationship between use of cohesive devices and writing quality has not been clearly shown. See Appendix C for overview of studies.

Of the five types of cohesion described by Halliday and Hasan in *Cohesion in English*, three have received the greatest attention in past studies: reference, conjunctive, and lexical cohesion. The following table provides a review of the types and their corresponding subtypes.

Table 1. *Cohesive devices analyzed in past research*

Reference	Conjunctive	Lexical
personal reference	additive	repetition
demonstratives	adversative	synonym
definite article	causal	superordinate
comparatives	temporal	general term
		collocation

Before turning to the details of the relevant studies, a related field of inquiry must be acknowledged: corpus-based research on overuse and misuse of connectors in L2 writing (Bolten et al, 2002; Granger and Tyson, 1996; Milton and Tsang, 1993). Although such research falls outside the scope this project, it is worthwhile noting that many of the studied connectors belong to Halliday and Hasan's category of conjunctive devices. In general, these studies suggest that L2 English writers overuse and misuse certain logical connectors with negative effects on their writing (*for instance, on the contrary, however, instead, therefore* etc). One study in the literature, Hinkel (2001), bridges the two types of research, using corpus-linguistics methodology and categories of cohesion based on Halliday and Hasan's framework to compare native and non-native writing.

Of the cohesion studies, only Witte and Faigley (1981) identify reference cohesion as an important individual factor. Compared with conjunctive and lexical cohesion, reference devices formed the second largest group after lexical devices in most studies and ranged from 10.8% to 29.8% of all cohesive devices (Zhang, 2000 and Liu and Braine, 2005 respectively). In detailed studies, pronominal reference devices constituted the largest group of reference devices ranging from 48.5% to 70.7% of total reference cohesion ("good" essays, Neuner, 1987 and "poor" essays, Neuner, 1987, respectively) . Demonstratives comprised the next largest group of reference devices, followed by definite articles and comparatives (Braine and Liu, 2005; Neuner, 1987; Zhang, 2000). According to Hinkel's 2001 analysis, non-native writers used significantly more demonstratives than their native speaking peers, often to the detriment of their writing.

In most studies, high-rated or advanced writing exhibited greater use of conjunctive devices than low-rated or intermediate writing (Johnson, 1992; Neuner, 1987; Tierney and Mosenthal, 1983; Wenjun, 1999, Witte and Faigley, 1981 but not Zhang, 2000). With the

1.3 Research Questions

Given the contradictory findings in past research, the question of cohesion and its relation to writing qualities merits further investigation. However, instead of following a comparative design (good versus poor writing or native versus non-native writing), this research project seeks to provide a more nuanced picture of cohesion in successful L2 writing and to investigate to what extent use of cohesive devices varies across successful L2 writers and across different writing tasks. Thus, the design, coding, and analysis of this research project was guided by the following overarching question and its two corresponding subquestions:

What is the relationship between writing quality and the use of cohesive devices in academic writing?

RQ1: Which cohesive devices do successful L2 writers use in their writing?

RQ2: In what ways does use of cohesive devices vary between successful L2 writers and across genres?

PART TWO: METHODOLOGY

2.1 Participants

Data for this study on the use of cohesive devices in academic L2 English came from the writings of four undergraduate students enrolled in the advanced writing course *ELI-100: Expository Writing: A Guided Approach* at the English Language Institute (ELI) at the University of Hawaii-Manoa (UH). The University of Hawaii's English Language Institute (ELI) is an academically oriented English language instruction department that provides courses for international and immigrant students admitted to the university in order to prepare them for their studies. At the time of enrollment in the L2 English writing course, all students were concurrently matriculated in degree-seeking programs at UH. **Students were required to take the writing course on the basis of their TOEFL scores (range = 500-600) and their placement test**

results and/or progression through the English Language Institute's preparatory writing courses.

Within the university, *Expository Writing: A Guided Approach* is considered comparable to an introductory freshman English course, and students receive letter grades for their performance.

Upon completion of this course, students finish the ELI's writing curriculum and continue their studies in their respective degree programs.

Data collection consisted of two stages: soliciting participants and selecting research candidates. Participants were solicited in two sections of *ELI-100: Expository Writing: A Guided Approach*, both taught by the same graduate instructor. Of the four sections of ELI-100 offered, these two sections were targeted because it was anticipated that the genres of the three major writing projects during the semester would challenge students with different writing demands and provide diverse opportunities for using cohesive devices. To address the project's research questions, four participants were chosen from a pool of eight volunteers based on their grades for the first two writing assignments. All four participants received A-'s or better on all three papers based on rubric grading schemes that assessed content, structure, and language. In a course where grades ranged from F to A, papers receiving grades of A- or better were considered examples of successful academic writing and therefore appropriate for investigating the frequency, use, and range of cohesive devices in successful L2 writing. To preserve anonymity, participants are referred to with pseudonyms.

2.2 Data

Data collected for the research project belong to three broad genres of academic writing: definitions, text analyses, and research papers. In the first assignment, students were asked to characterize and define the construct *good writing* using supporting examples (Appendix D). For the second assignment, students completed a careful text analysis while addressing the question

Kenton Harsch 6/18/07 8:41 AM

Comment: In fact, TOEFL is not used for determining undergraduates' writing needs. The Writing Placement Exam (Manoa Writing Program) is used for native and non-native writers of English alike, and placement is based on whether their writing appears to be native-like (English Department courses) or non-native-like (ELI).

of reading improves writing skills (Appendix E). The third and final assignment required students to undertake an in-depth library research project on a topic of their choice, culminating in an extended research paper (Appendix F). All writing samples are final drafts. Students revised their early drafts after receiving instructor feedback and completing editing and revising exercises. Students submitted their data to the researcher electronically via e-mail.

Although the assignments varied in length depending on the writer, the definition paper was the shortest ($M = 867.25$ tokens) across the participants, followed by the text analysis ($M = 1550.25$ tokens), and the research paper ($M = 2299.75$ tokens). Table 2 provides raw numbers for each participant's data as well as the average length of each text type, the total token counts for each task, and the overall size of the corpus. Orthographic words served as the basis for token counts, and paper titles were included while reference lists and block quotes were excluded. The total size of the corpus was 18,869 tokens.

Table 3. *Corpus size by individual and task type*

Text Type	Participants				Tokens/Text	Totals
	Marika	Saori	Quynh	Yukiko		
Definition Paper	879	1247	676	667	867.25	3469
Text Analysis Paper	1349	1205	1675	1972	1550.25	6201
Research Paper	1710	1813	2488	3188	2299.75	9199
Totals	3938	4265	4839	5827		18869

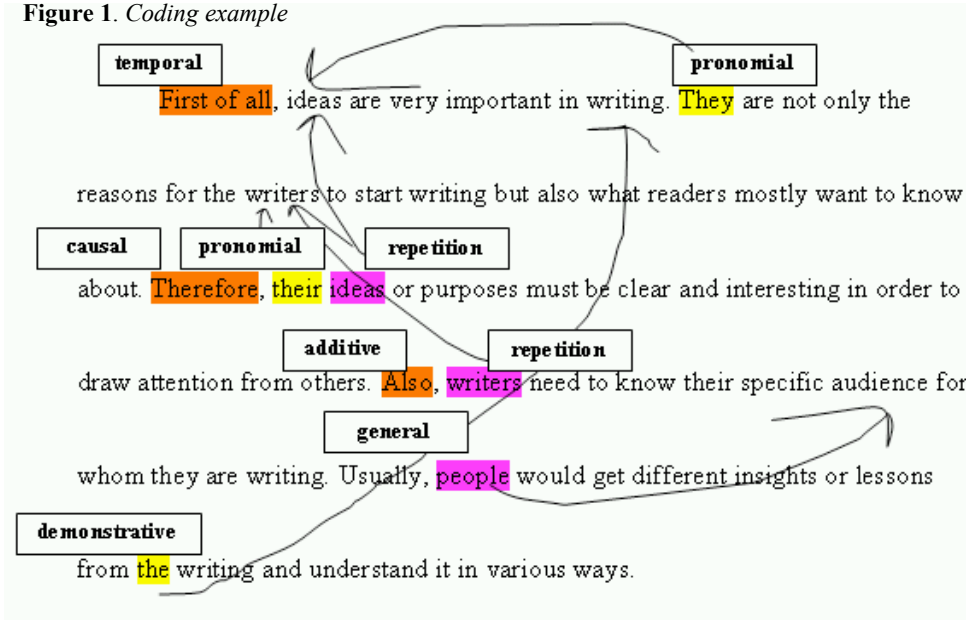
2.3 Coding

Coding followed Halliday and Hasan's taxonomy for cohesive devices, focusing on reference, conjunctive, and lexical devices and their corresponding subtypes. As in previous studies, substitute and ellipsis were excluded from the coding and analysis. During the first coding phase, each writing sample was checked manually for reference, conjunction and lexical

cohesion, one type at a time. While coding for reference, cohesive items and their corresponding referents were identified. Similarly, lexical cohesion was traced by creating chains of repetitions, synonyms, and superordinates across sentences and paragraphs.

During the second coding phase, previously identified cohesive items were double-checked and recorded in an Excel database, searchable by participant, genre, cohesive device type and subtype. To minimize ambiguity and insure consistency in the data coding, an item was considered cohesive if and only if it crossed sentence boundaries. This followed a definition of *cohesive tie* proposed by Halliday and Hasan (1976, p. 329) “a single tie between a pair of elements in adjacent [or more remotely separated] sentences.” It is important to note that some previous cohesion studies also considered sentence internal ties when calculating cohesion (e.g., Braine and Liu, 2005).

Figure 1. Coding example



In the coding example above, orange indicates conjunctive devices, yellow marks reference conjunction, and pink shows lexical cohesion. For each color coding (cohesive device type) a cohesive device subtype was specified with a label. For example, *First of all*, colored orange for conjunctive device, has been coded with the label *temporal* for its cohesive subtype.

Although each type of cohesion presented different coding problems, reference and lexical proved significantly more challenging than conjunction. Locating the referents of demonstrative cohesive ties was difficult for two reasons, illustrated by the examples below: 1) some writers used demonstrative cohesive device, such as *the*, without prior mention of the referent and 2) the demonstrative head and modifier *this* was used both endophorically (referring to a previously mentioned element in the text, i.e., cohesively) and exophorically (referring to an external entity existing outside the body of text, i.e., not cohesively). Cases of missing referents and exophoric reference were not counted as examples of cohesion and were not included in the following analyses.

- 1) My five senses...were in operation as if I sat in *the* 3 dimensional movie theater.
- 2) So, students in another major or student [sic] not interested in human development may think *this* text is not as interesting as I think it is.

With regard to lexical cohesion, the researcher used her best judgment to determine synonyms and subordinate/superordinate terms, albeit it must be acknowledged that these codings are subjective at best. Finally, due to time constraints, lexical cohesion coding was conducted for Tasks 1 and 2 only.

PART THREE: RESULTS AND ANALYSIS

To provide a detailed portrait of cohesion in successful L2 academic writing, the following results are presented and discussed from three different perspectives: overall cohesion, cohesion by individual, and cohesion by task type.

3.1 Overall cohesion

Broadly viewed, this study's results corroborate previous findings in cohesion studies regarding the frequency and distribution of cohesion types and subtypes. As illustrated in Table 3 below, lexical cohesion comprised the largest percentage of total cohesive ties, ranging from 61% for Task 1 to 64% for Task 2. Following lexical ties, reference ties accounted for 22% (Task 1) to 25% (Task 2) of cohesion. Conjunctive devices made up the smallest percentage of ties, 14% for both writing tasks. Although lexical cohesion for Task 3 was not analyzed and no overall distributions for cohesion types can be calculated, findings for Task 1 and Task 2 are consistent with each other and with previous research studies (see Section 1, pp. 6-8).

Table 4. Overview of cohesion types across three writing tasks

	Reference		Conjunction		Lexical		Total
	N	%	N	%	N	%	N
Task 1	98	25%	57	14%	241	61%	396
Task 2	126	22%	84	14%	376	64%	586
Task 3	91	-	104	-	-	-	195 ^a
Total	315	-	245	-	617^a	-	1177^a

^aindicates that lexical cohesion counts for Task 3 were not included

Unlike general findings for the main cohesion types, results for specific reference subtypes in this study differed from previous studies. Whereas early studies found pronomial reference to be the largest subtype of reference cohesion (Braine and Liu, 2005; Neuner, 1987; Zhang, 2000), coding in this study revealed demonstrative ties as the most frequent subtype, comprising 49% of all reference ties. Pronomial ties, the second largest group, made up 45% and comparative ties accounted for the remaining 6% of reference cohesion. Based on figures in Table 5, one can see that the five most frequent ties *it* and *they* (pronomial) and *this*, *the*, and *these* (demonstrative) covered 78% of all reference cohesion together.

Table 5. Subtypes and frequencies of reference devices

Pronominal		Demonstrative		Comparative	
tie	frequency	tie	frequency	tie	frequency
it	51	this	97	another	4
they	45	the	28	same	4
he	19	these	25	more	3
their	10	those	3	such	2
his	9	that	2	different	2
them	7			each	1
she	1			other	1
				similar	1
	142		155		18
	45%		49%		6%

The following two examples illustrate pronominal, demonstrative, and comparative reference ties in the data.

- (1) What is good writing? To answer **this** ^[demonstrative; referent: *what is good writing?*] question, readers can probably tell **it** ^[pronominal; referent: *good writing*] right away... (Quynh, Task 2)
- (2) So, as you read on, please compare your beliefs on “good writing” with mine. Are **they** ^[pronominal; referent: beliefs] **the same** ^[comparative; referent: your beliefs] or are **they** ^[pronominal; referent: beliefs] **different** ^[comparative; referent: your beliefs]? (Marika, Task 2)

Results for conjunctive subtypes correspond with general patterns in previous findings with additive devices as most frequent (32%), followed by temporal (30%), adversative (22%), and causal (16%) devices. Of the additive devices, *also*, *and*, *for example*, and *in addition* were most frequent, constituting 73% of all additive ties. Among adversative ties, *however* and *but* comprised 64% of adversative cohesion. *So*, *therefore*, and *as a result* were the most common causal ties (65%). Finally, the text-structuring temporal devices *first(ly)*, *second(ly)*, *third(ly)*, and *finally* comprised 59% of temporal cohesion. Of the four different conjunctive subtypes, temporal showed the greatest variety and included 22 different cohesive ties, even when collapsing *first* and *firstly*, *second* and *secondly* etc., together. A full list of conjunctive ties follows in Table 6.

Table 6. Subtypes and frequencies of conjunctive devices

Additive		Adversative		Causal		Temporal	
tie	freq	tie	freq	tie	freq	tie	freq
also	20	however	23	so	11	first(ly)	16
and	12	but	12	therefore	9	second(ly)	10
for example	12	on the other hand	7	as a result	6	third(ly)	9
in addition	12	in fact	6	consequently	2	finally	8
furthermore	5	on the contrary	2	hence	2	succeeding	5
moreover	5	as a matter of fact	2	on the basis	2	last	3
for instance	4	nevertheless	2	then	2	then	3
in other words	3	in contrast	1	to that end	2	after +	2
besides	2			as a consequence	1	following	2
likewise	2			from this	1	fourth	2
				in this case	1	in summary	2
				thus	1	three	2
						above	1
						at the same time	1
						in conclusion	1
						next	1
						one	1
						previous	1
						tertiary	1
						to begin with	1
						under these circumstances	1
	77	55	40	73			
	32%	22%	16%	30%			

The following four examples illustrate students' use of the four conjunctive subtypes:

- (3) Additive:
It gains a momentum one paragraph after another, so I have no time to get bored.
Moreover...(Saori, Task 1)
- (4) Adversative:
Eventually there are six traits of writing that we have talked in class. For me,
however, the most important things...(Quynh, Task 1)
- (5) Causal:
So, students in another major or student not interested in human development may think this text is not as interesting...(Marika, Task 2)
- (6) Temporal:
To begin with, I summarize the chapter titled...(Saori, Task 2)

Finally, results for lexical subtypes matched previous studies with repetition as the largest group (74%), followed by synonyms (22%), superordinates (4%), and general words (2%).

Below, Table 7 provides an overview of the raw frequencies for each subtype.

Table 7. *Subtypes and frequencies of lexical devices*

Repetition		Synonym		Superordinate		General		Total
N	%	N	%	N	%	N	%	N
476	74	110	22%	22	4%	9	2%	617

As in the discussion of reference and conjunctive subtypes above, examples to illustrate participants' use of lexical subtypes are also provided. Although the example of superordinate lexical cohesion in (9) shows cohesion across different parts of speech, verbs (*smell, feel, hear*) and a noun (*senses*), it has been included in analysis because it is a much more original and creative example than many of the other instances (e.g., *textbook/text, visual arts/arts, writing/piece of work*).

- (7) Repetition:
What is **good writing**? What makes a **writing** piece **good**? (Marika, Task 1)
- (8) Synonym:
Thirdly, as in screenplay, the **structure** of the scenes are [sic] very important [intervening sentence]...similar to the **composition** of any text (Quyhn, Task 2).
- (9) Superordinate:
While I was reading, I was almost able to **smell** cider trees, **feel** the temperature of the air at the time, and **hear** the sounds of cicadas. My five **senses**...(Saori, Task 1).
- (10) General word:
Also writers need to know their specific **audience** for whom they are writing. Usually **people** would get different lessons or insights from the writing (Quyhn, Task 1).

3.2 Cohesion by individual

To investigate variation in cohesion across individual writers, raw frequency data was normed to account for text length in two ways. As summarized in Table 2 (p. 11), total individual

production varied widely across writers with a maximum of 5827 tokens (Yukiko) and a minimum of 3938 tokens (Marika), creating a range of 1889 tokens between different participants' production. The first adjustment method addressed the issue of text length by calculating each participant's cohesion for each type of cohesion per 1000 tokens, as shown in Table 8. In the lexical cohesion column Table 8 only Task 1 and Task 2 were included in the norming; lexical cohesion for Task 3 was not calculated. The second adjustment method incorporated the definition of cohesive device as a sentence crossing tie and calculated cohesion per sentence, illustrated in Table 9. Only cohesive devices and sentences in Tasks 1 and 2 are included in Table 9.

Table 8. *Types of cohesion and normed frequency counts by individual (per 1000 tokens)*

Individual	Reference		Conjunctive		Lexical	
	N	N/1000	N	N/1000	N	N/1000
Marika	59	14.98	47	8.07	165	74.06
Saori	53	12.43	49	11.49	155	63.21
Quynh	67	13.85	93	19.22	140	59.55
Yukiko	136	23.34	57	9.78	157	59.49

Table 9. *Cohesion per sentence by individual for Task 1 and Task 2*

Individual	Cohesive Devices (N)	Sentences (N)	Cohesive Devices/Sentence
Marika	234	111	2.11
Saori	227	143	1.59
Quynh	231	100	2.31
Yukiko	290	145	2.00

Both Tables 8 and 9 show sometimes uniform, sometimes varying use of cohesive devices across individual writers. Table 8 reveals that use of cohesion varied little across participants for reference and conjunctive devices and diverged considerably across participants for lexical devices. With the exception of Yukiko (23.34 reference ties/1000 tokens), participants used between 12.43 and 14.98 reference ties per 1000 tokens. Similarly, only one student, Quynh (19.22 conjunctive ties/1000 tokens), employed noticeably more conjunctive devices than the

other participants, who ranged from 8.07 to 11.49 conjunctive devices per 1000 tokens. Normed lexical device figures, however, create a murkier picture of cohesion; participants ranged from 59.49 to 74.06 lexical devices per 1000 tokens. Table 9 also shows a distribution of participants across a range of 1.59 cohesive ties per sentence to 2.31 cohesive ties per sentence.

3.2 Cohesion by task type

To analyze variation in cohesion across text types, raw frequencies of cohesive devices were normed per 1000 tokens for each of the different writing tasks. The results, as summarized in Table 10, show a marked decrease in the use of all cohesive devices from Task 1 to Task 2 to Task 3. Although lexical cohesion counts for Task 3 were not calculated, the available results for reference, conjunction, and lexical devices suggest that a similar decrease in lexical cohesion could be anticipated for Task 3.

Table 10. *Types and frequency of cohesion by task*

	Reference		Conjunction		Lexical	
	N	N/1000	N	N/1000	N	N/1000
Task 1	98	28.25	57	16.43	241	69.47
Task 2	126	20.32	84	13.55	376	60.64
Task 3	91	9.89	104	11.31	-	-
Total	315	16.69	245	12.98	617	63.81

PART FOUR: DISCUSSION AND CONCLUSION

Inspired by contradictory findings and an overwhelming focus on comparing rather than describing cohesion in writing in previous research, this study aimed at providing a comprehensive landscape of cohesion in academic L2 writing. By focusing on “successful” academic writing, this research sought to outline and analyze the types and frequencies of

cohesion types in L2 writing that is judged to be “good” in an academic setting. To accomplish this goal, coding and analysis followed two guiding research questions: 1) Which cohesive devices do successful L2 writers use in their writing? and 2) In what ways does use of cohesive devices vary between successful L2 writers and across genres?

In response to the first research question, analysis revealed that successful L2 writers employed all three types of cohesive devices as well as the many subtypes of cohesion. Lexical devices comprised the largest share of cohesion, followed by reference and pronomial, as found repeatedly in other studies of L1 and L2 writing. Within lexical cohesion, repetition dominated, accounting for more than three-quarters of all lexical devices. Such findings hold true across past studies and suggest, as noted by Witte and Faigley (1981), that increased use of synonyms may characterize very advanced writers, labeled “high-rated writers” in their study. Unlike past studies on cohesion, demonstratives in this study made up the largest percentage of reference devices, followed closely by pronomial reference and limited use of comparative reference. This finding, although at odds with earlier studies, mirrors Hinkel (2001), which showed that non-native writers used significantly more demonstrative pronouns than native-speaking peers, often in ways that made their texts “confusing” to the reader (p. 124). Although only a few instances of demonstrative cohesion proved incomprehensible, locating the appropriate referents for reference devices, particularly demonstrative and pronomial ties, was challenging at times. Finally, within the conjunctive type, additives and temporals exceeded adversatives and causals. For the most part, the most frequent items matched findings in previous studies.

Results for the second two-fold research question seemed surprising, both regarding the variation between learners and the variation across writing tasks. With the exception of one participant for each type, reference and conjunctive cohesion use was unexpectedly uniform

across writers. Lexical cohesion use, on the other hand, differed more drastically between writers. When number of cohesive ties per sentence was calculated for each individual, interpretation of the results posed a conceptual problem. What does it mean if one participant uses 1.59 cohesive ties per sentence while another deploys 2.31? In a text comprising 50 sentences, such rates amount to a difference of 36 cohesive devices between the writers. Thus, the most cautious interpretation appears to be that some variation existed between the advanced L2 writers in this study on different levels of cohesion use (cohesive ties per sentence, lexical cohesion), but their overall use of cohesive devices was similar. Variation in cohesive devices across text types, in contrast, was repeatedly and noticeably present in the data. Although relative frequencies of cohesive devices remained constant across task types, cohesion frequency rates for all three types decreased consistently from Task 1 to Task 2 to Task 3. Thus, though the identity of the writer did not always influence frequency of cohesive devices, the type of text the writer wrote repeatedly did affect the amount of cohesion used.

A possible explanation for the decrease in cohesion density from Task 1 to Task 3 is text length. Whereas Task 1 averaged 867.25 tokens, Task 3 averaged 2299.75 tokens, representing an increase of over 250% in text length from the first to the third assignment. Perhaps in shorter texts, writers feel obligated to write shorter, more compact and cohesive texts, full of strong links between sentences and paragraphs. However, given the drastic decrease in reference cohesion, from 28.25 to 9.89 references devices per 1000 words, one would expect to see some sort of corollary increase in other overt forms of cohesion, such as repetition or synonymy. That is, instead of using a pronomial to refer to an aforementioned element, the writer repeats or paraphrases the term. However, this expectation is not borne out in the lexical data, at least not for Task 2. Possibly tracing the different reference subtypes across tasks could shed further light

on this mystery. If demonstrative ties account for a large portion of reference cohesion in Tasks 1 and 2 but decreases dramatically in Task 3, such a decrease could account for part of the overall decline in reference cohesion. Given these inconclusive postulations, an intriguing and important area of future research would be exploring the influence of text type and text length on cohesion in greater depth.

When reflecting on the findings of this project, several important issues came to the forefront. First, are these text types truly diverse enough represent different genres? Although text length may play a role in amount of cohesion, the relative frequencies of cohesion types remained the same across the three tasks. Thus, another area for research could be describing cohesion in more distinct genres, such as narrative prose, creative writing, and argumentative essays. Secondly, Halliday and Hasan's taxonomy does not include a category for morphologically related words (i.e., words from the same family or derived from the same stem such as *writing* (n)/*write* (v)/*writer* (n)). For evaluating lexical cohesion, the lack of an appropriate category for such related words means writers receive no credit for creating cohesive strings of text and demonstrating their knowledge of word families, as in the following examples:

- (1) His first sentence introduces a main idea with *clarity*...[intervening sentences]...In this way, his main idea becomes crystal *clear*...(Saori, Task 2).
- (2) ...[Writing] requires a lot of efforts from writers in using language to depict or impress artists' mood and feeling and the composition in each artwork so that readers can easily *imagine* or capture its meanings...[intervening sentences]...[good word choice] also create[s] an *image* in their mind (Quynh, Task 2).
- (3) In the text, Feldman talks about the physical growth the nervous system and brain, and the life cycle of *infancy*. The chapter talks about the rate of growth, the four principles of growth, the environmental influences, and the rhythm or sleep patterns *infants* have (Marika, Task 2).

- (4) After I finished *reading* this article, I realized how excited I was while I was *reading*. I could not wait to finish *reading* it. The author ends each paragraph with a sentence that makes *readers* wonder what will be in the next paragraph (Yukiko, Task 2).

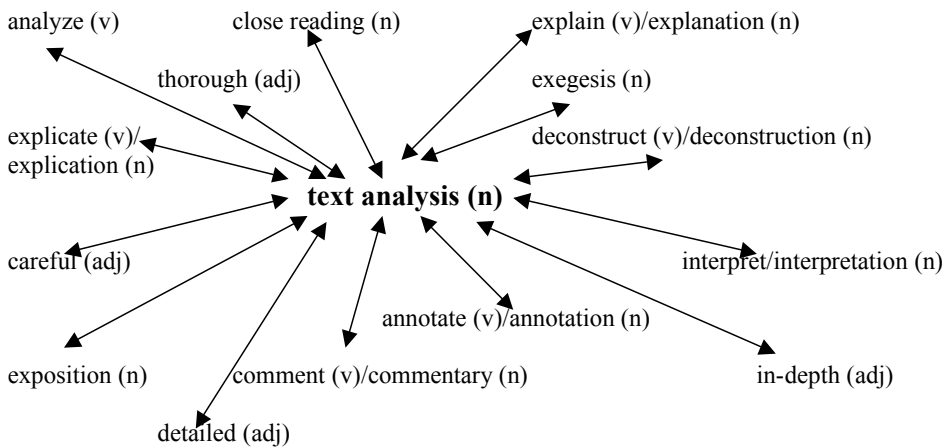
To provide an accurate picture of overall lexical cohesion, instances like the examples above must be incorporated in some way, perhaps as an additional subtype.

A third and final issue has been raised by several researchers and enlivened debate over the usefulness of cohesion studies. In essence, researchers such as Carrell (1982) have argued that cohesion is not coherence; that is, the presence of cohesive devices does not guarantee a comprehensible and coherent text. This study acknowledges that over-reliance on cohesive devices detracts from writing quality rather than promoting it, particularly excessive use of the conjunctive devices identified in corpus studies (Bolten et al, 2002; Granger and Tyson, 1996; Milton and Tsang, 1993). For this reason, this study took successful L2 writing as its departure point with the goal of providing an in-depth description of cohesion in successful writing rather than correlating use of particular cohesive devices with the constructs of good or bad writing.

Before closing, several broad pedagogical implications merit discussion on the basis of these research findings. Although these writings represent the work of already successful L2 academic writers, two possible areas for future instruction emerge. First, in line with this study's results and Hinkel (2001)'s findings, L2 writers appear to need additional instruction on the use of pronomial and demonstrative reference devices. Such guidance could take two forms: (1) (re)inform L2 writers that beginning sentences with pronomials or demonstratives as heads is often regarded as "weak" writing. These often vague pronomials and demonstratives are best replaced with more descriptive noun phrases, either repetitions or synonyms of the referent element (see Examples 3, 4) and (2) (re)remind L2 writers to check and ensure that referents for demonstratives and pronomials are clear (5).

- (3) *Rape* ...[...] *It* is the action...[...] *It* is one of the cruelest crimes (Yukiko, Task 3).
- (4) But one thing is true and that is, not all readers are going to feel an attraction to this novel. *This* is because each individual sees good writing in a different way...(Marika, Task 1).
- (5) I also included sentence fluency and clarity because if the writing is not clear or fluent, the readers are not going to understand the main idea or topic of the writing. And *this* is made difficult, as there are many different ways to view a piece of work (Marika, Task 1) .

The second pedagogical implication is the need for lexical diversity. Although the participants in this study are accomplished and advanced L2 writers, their writing would benefit from an emphasis on expanding their vocabularies. Rather than relying on repetition alone, all writers, L1 and L2, must develop a wide range of synonyms and superordinates, which allows for variety in writing. Two possible approaches for encouraging practical and grounded vocabulary development include 1) habitual yet judicious use of a good thesaurus while writing and 2) on-going construction of personal vocabulary matrices. For the former, teachers should focus on helping L2 writers hone their abilities to select context-appropriate synonyms from a thesaurus. For the latter, instructors should encourage students to create vocabulary matrices for topics as a pre-writing activity. For example, an model vocabulary matrix for Task 2, the text analysis assignment, might resemble the following:



To conclude, this pilot study of cohesion in L2 academic writing has attempted to paint a detailed portrait of cohesion in successful L2 academic writing and provide suggestions for teaching writing in the classroom. Despite numerous limitations, such as number of participants, lack of interrater reliability on codings, restricted variety of genres, lacking access to students' original drafts during the writing process and a host of other issues, this research suggests that the study of cohesion in academic writing is indeed of merit, both to the research and the teaching communities. Although no broad generalizations can be derived from this study, the findings suggest cohesion may differ more across text types than across individual writers. To evaluate this hypothesis, careful future research is needed both across different L2 writers and across different writing tasks. Lastly, in response to the paper's questioning title, successful L2 writers DO use a wide variety of cohesive devices, but as is always true with the evolving art of writing, there is always room for improvement.

PART FIVE: BIBLIOGRAPHY

- Bolton, K. et al. (2002). A corpus-based study of connectors in student writing. *International Journal of Corpus Linguistics*, 7(2), 165-182.
- Carrell, P. (1983). Cohesion is not coherence. *TESOL Quarterly*, 16(4), 479-488.
- Granger, S. & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17-27.
- Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hinkel, E. (2001). Matters of cohesion in L2 academic texts. *Applied Language Learning*, 12(2), 111-132.
- Johnson, P. (1983). Cohesion and coherence in compositions in Malay and English. *RELC*, 23(2), 1-17.
- Liu, M. & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System*, 33, 623-636.
- Milton, J. & Tsang, E.S. (1993). A corpus-based study of logical connectors in EFL students' writing: directions for future research. In R. Pemberton & E. Tsang (Eds) *Studies in Lexis*, 215 – 246. Hong Kong: HKUST.
- Tierney, R. & Mosenthal, J. (1983). Cohesion and textual coherence. *Research in the Teaching of English*, 17(3), 215-229.
- Wenjun, J. (1999). A quantitative study of cohesion in Chinese graduate students' writing: Variations across genres and proficiency levels. Paper presented at the Symposium on Second Language Writing at Purdue University.
- Witte, S. & Faigley, L (1981). Coherence, cohesion, and writing quality. *College composition and communication*, 32(2), 189-204.
- Zhang, M. (2000). Cohesive features in exploratory writing of undergraduates in two Chinese universities. *RELC*, 31(1), 61-93.

PART SIX: APPENDICES

Appendix A: Overview of reference devices

Personal Reference				Demonstrative Reference				Comparative Reference		
pronoun	determiner			determiner	adverb			determiner	adjective	adverb
<i>I</i>	<i>me</i>	<i>mine</i>	<i>my</i>	<i>this</i>	<i>these</i>	<i>here</i>	<i>now</i>	<i>the</i>	<i>same</i>	<i>identically</i>
<i>you</i>		<i>yours</i>	<i>your</i>	<i>that</i>	<i>those</i>	<i>there</i>	<i>then</i>		<i>identical</i>	<i>similarly</i>
<i>we</i>	<i>us</i>	<i>ours</i>	<i>our</i>						<i>equal</i>	<i>likewise</i>
<i>he</i>	<i>him</i>	<i>his</i>	<i>his</i>						<i>similar</i>	<i>so</i>
<i>she</i>	<i>her</i>	<i>hers</i>	<i>her</i>						<i>additional</i>	<i>such</i>
<i>they</i>	<i>them</i>	<i>their</i>	<i>their</i>						<i>other</i>	<i>differently</i>
<i>it</i>		<i>s</i>	<i>its</i>						<i>different</i>	<i>otherwise</i>
<i>one</i>		<i>its</i>	<i>one's</i>						<i>else</i>	<i>so</i>
									<i>better</i>	<i>more less</i>
									<i>more</i>	<i>equally</i>